

# Automated Scenario Generation for Meeting Human-in-the-Loop Simulation Requirements

Gano B. Chatterji\*

Crown Consulting, Inc., Moffett Field, CA, 94035-1000

and

Yun Zheng†

Crown Consulting, Inc., Moffett Field, CA, 94035-1000

**A two-step automated Multi-Aircraft Control System traffic scenario generation process for Human-in-the-Loop evaluations of air traffic management concepts is described. The first step of the two-step process employs the scenario generation capability currently available in NASA's Air Traffic Management Testbed. The second step refines the scenario by filtering flights from the traffic scenario based on route length, cruise speed, cruise altitude, entry time and the desired ratio of internal to external flights. A solution for achieving the desired ratio of internal to external flights, where internal flights are shorter flights and external flights are longer flights based on a distance threshold, is described. Finally, schedulers are described for shaping the hourly arrival traffic count as a function of time in response to airport capacity constraint or for increasing the traffic demand with respect to the available arrival capacity. Results generated for arrival traffic to the four major airports in the New York Metroplex on a busy day using the two-step procedure are discussed. These results show that traffic scenarios for Multi-Aircraft Control System that meet the Human-in-the-Loop and fast-time simulation requirements can be created automatically following the procedures described in the paper. The automated process will improve the accuracy and efficiency by eliminating the tedious manual process for scenario generation.**

## I. Introduction

The automated scenario generation capability in the Air Traffic Management (ATM) Testbed (ATMTB) has been used to generate Multi-Aircraft Control System (MACS) scenarios for Human-in-the-Loop (HITL) evaluations. MACS is a distributed simulation system with multiple pseudo-pilot and air traffic controller stations that is frequently used for National Aeronautics and Space Administration (NASA)'s evaluations of air traffic management concepts.<sup>1</sup> Automated scenario generation has been used for creating MACS scenarios for Dynamic Routes for Arrivals in Weather (DRAW)<sup>2</sup> and Integrated Demand Management<sup>3</sup> HITL experiments. It has also been used for creating Airspace Target Generator (ATG) scenarios, used in realistic airport surface traffic simulation, for the Airspace Technology Demonstration (ATD-2)<sup>4</sup> subproject. Automated scenario generation is currently being used to generate MACS scenarios for Instrument Flight Rules (IFR) traffic, Visual Flight Rules (VFR) traffic and the expected Urban Air Mobility (UAM) traffic for evaluations under the ATM-eXploration (ATM-X) project<sup>5</sup> to enable simulation of future UAM vehicles to operate in the National Airspace System (NAS).

This paper describes a two-step automated process for creating MACS traffic scenarios according to the desired scenario characteristics. The first step of the two-step procedure described in this paper employs the ATMTB automated scenario generation process first described in Ref. 6. The second step, introduced in this paper, enhances the scenario, created by the first step, by eliminating flights with unreasonable values associated with parameters such as cruise speed and cruise altitude, and selecting/adjusting flights based on desired scenario characteristics specified by the user such as route length, entry time, landing rate and ratio of number of flights inside to outside the terminal area (for example with respect to the Time-Based Flow Management freeze horizon).

Prior to the automated scenario generation capability in ATMTB, traffic scenarios for MACS were generated manually by first creating an initial scenario (seed-scenario) by selecting flight-plans from recorded air traffic data

\* Senior Scientist and Lead, Crown Consulting, Inc., MS 210-8, Associate Fellow.

† Senior Software Engineer, Crown Consulting, Inc., MS 210-8.

and then modifying the scenario, running it in MACS, modifying the scenario based on the results, and repeating the process until the characteristics desired for meeting the objectives of the HITL test were achieved. Understandably, this process is time consuming. Even creating a seed-scenario that results in successful MACS simulation was difficult because of missing and erroneous data. With the objective of replacing the difficult manual process with the automated process, the earlier research described in Ref. 6 used the ATMTB automated scenario generation process to create a seed-scenario that served as a starting point for further manual adjustments for creating the HITL-scenario. Analyses were described in that study for comparing the seed and the HITL-scenarios for identifying data quality issues and for eliminating flights with unreasonable parameter values from the seed-scenario. Comparison of the seed-scenario with the manually altered HITL-scenario revealed that many of the desired characteristics, such as flights having unique call-signs and airport arrival demand exceeding the airport arrival capacity, can also be achieved by enhancing the automated scenario generation process. This finding is the inspiration for the second step of the two-step process.

The motivation for the work described in this paper is the need for efficiently creating scenarios for HITL and fast-time evaluations of Trajectory Based Operations (TBO) concepts like the Integrated Demand Management (IDM). This concept seeks to integrate capabilities of traffic flow management and terminal-area traffic management to organize aircraft trajectories into well-managed flows that match traffic demand to the available capacity especially into and out of airports in the New York Metroplex. The automated process will also enable creation of a large number of scenarios needed for Monte-Carlo runs that is difficult to accomplish manually.

The rest of the paper is organized as follows. With the focus on MACS traffic scenarios for investigations of operational feasibility of the TBO concepts, the IDM concept is briefly described in Section II. Given the attention of these projects on applications to the Northeast region of the NAS, results of analysis of hourly runway configurations for identifying days with high-volume of arrival and departure traffic at four airports- John F. Kennedy (JFK), Newark Liberty (EWR), LaGuardia (LGA) and Teterboro (TEB), for every day of 2017, are summarized in Section III. This includes tables with the most frequently used runway configurations at each of the four airports individually and together, and hours on days with a large number of arrival and departure operations. Based on this analysis, traffic data from 5/23/2017 were used for creating the MACS scenario. The first step of two-step process, which is the ATMTB scenario generation process, is discussed in Section IV. The second step of removing outliers and conditioning the scenario is described in Section V. Algorithms are described for achieving desired scenario characteristics and results are presented in this section to demonstrate the ability of the two-step automated scenario generation process to create scenarios that meet the scenario requirements for HITL and fast-time evaluations of IDM concept using MACS. Finally, the main findings are summarized in Section VI.

## **II. Background**

The research focus for IDM<sup>7</sup> Trajectory Based Operations (TBO) concepts is improving the efficiency and predictability of air traffic operations. Many of the air traffic management tools and technologies in the enroute, terminal-area and surface domains developed both by the Federal Aviation Administration (FAA) and NASA that are used for operations in the National Airspace System have been difficult to use in the congested Northeast region airspace due to a mix of high traffic-volume, weather conditions, the proximity of major airports in the New York Metroplex and in neighboring centers, airspace geometry, and operational procedures for separating the flows in and out of major airports. TBO concepts seek to collaboratively organize aircraft trajectories into well-managed flows that match traffic demand to the available capacity by initially leveraging FAA and NASA pre-departure, enroute, arrival-departure and surface technologies.<sup>8</sup>

The charter of the IDM project is to explore ways of integrating near-term to mid-term NextGen traffic management systems to improve efficiency in situations when the traffic demand exceeds capacity of resources such as airports and airspace. In the IDM concept, FAA's Traffic Flow Management System (TFMS) and the Time-Based Flow Management (TBFM) system are used.<sup>7</sup> In the current system, TFMS is used strategically for determining departure delays at the airports of origin in response to constraints at destination airports and flow constrained areas (FCA). TBFM is a terminal-area traffic management system that tactically assigns a scheduled time of arrival (STA) to arrivals at the metering locations such as meter-fixes, meter-arcs and runway threshold based on capacity constraints at those locations. TFMS uses a Ration-By-Schedule algorithm that is based on the STA filed by the aircraft operator while TBFM computes an estimated time of arrival (ETA) using track and flight-plan data for its STA assignment. IDM seeks to establish coordination between strategic TFMS and tactical TBFM decisions for reducing delays by using TFMS to precondition traffic into the airspace domain of the TBFM system. Unfortunately, incorrect capacity forecast, delayed departure from the airport, wind and weather introduce uncertainty to the arrival time forecast, which disrupts the schedule and sequence intended by preconditioning. TBFM then has to impose additional delays to adjust the schedule for complying with the capacity constraints at the metering locations. Given that the uncertainty is higher

and the cost of delay is lower when the aircraft are on the ground compared to when they are airborne and close to the TBFM freeze-horizon boundary, a proper balance between TFMS and TBFM delays is needed for reducing fuel consumption (by delaying as little as possible while airborne), maintaining the airline schedule and fully utilizing the available airport capacity.

Several HITL and Automation-In-The-Loop experiments have been completed to investigate the operational feasibility of the IDM concept under realistic conditions.<sup>9, 10</sup> Recently, fast-time Monte-Carlo simulations are also being developed for IDM concept evaluations.<sup>11</sup> In these experiments, MACS simulates air traffic data based on the input traffic and weather/wind scenario files; it also provides a high-fidelity air traffic control simulation environment for controller and pilot interactions.

The progress made on the IDM project will be continued for developing a concept of operations and accompanying system architecture that evaluates the integration of FAA's systems and NASA's Airspace Technology Demonstration (ATD) ground-based and airborne systems for a future service-oriented airspace system.<sup>8</sup> This concept of operations needs to include the operation of new entrants such as supersonic aircraft, space launch vehicles, high-altitude long endurance platforms, Unmanned Aerial Systems and Urban Air Mobility vehicles in addition to the traditional airspace users like airlines and general aviation aircraft. The concept of operations developed is expected to be evaluated in a HITL simulation in the future using Northeast region scenarios developed with service provider and user inputs.

### III. Traffic Scenario Selection

This section describes the analysis of runway configurations at JFK, EWR, LGA and TEB with the objective of identifying days with high-volume of arrival and departure traffic for creating MACS traffic scenarios for TBO studies.

Because traffic flow patterns on the airport surface and in the terminal airspace depend on the runway configuration (runways used for arrivals and runways used for departures), hourly arrival and departure data for every day in 2017 were obtained for JFK, EWR, LGA and TEB airports from the FAA's Aviation System Performance Metrics (ASPM) database using the information in the Throughput Analysis Standard Report. These data were processed to determine the total number of hours each configuration was in use and the total number of operations (sum of arrivals and departures) conducted in each configuration during 2017. The top five configurations based on the percentage of hours in use are summarized in Table 1. The second column in the table shows the arrival runways such as 31L and 31R and departure runways such as 31L separated by a vertical bar (see the first row of the second column in Table 1). The

**Table 1 Top five most frequently used configurations at JFK, EWR, LGA and TEB**

Airport	Configuration	Hours	Arrivals	Departures	Operations	% Hours	% Operations
JFK	31L, 31R   31L	3,817	102,889	95,443	198,332	43	44
JFK	13L   13R	1,020	26,729	28,297	55,026	11	12
JFK	13L, 22L   13R	741	24,978	22,643	47,621	8	10
JFK	22L, 22R   22R	690	15,817	15,903	31,720	7	7
JFK	4L, 4R   4L	493	10,354	10,095	20,449	5	4
EWR	22L   22R	3,582	94,374	97,409	191,783	40	44
EWR	4R   4L	2,914	72,916	76,997	149,913	33	34
EWR	11, 22L   22R	496	18,856	16,270	35,126	5	8
EWR	22R   22R	334	2,673	1,786	4,459	3	1
EWR	4L   4L	287	2,501	1,891	4,392	3	1
LGA	22   13	2,041	46,927	46,222	93,149	24	25
LGA	31   4	1,519	36,906	36,394	73,300	18	20
LGA	4   13	1,385	32,576	33,406	65,982	16	18
LGA	22   31	1,210	28,715	28,706	57,421	14	16
LGA	31   31	1,175	21,098	21,294	42,392	14	11
TEB	19   24	4,050	39,002	37,752	76,754	47	48
TEB	6   1	3,001	25,575	25,665	51,240	35	32
TEB	19, 24   24	644	7,403	7,943	15,346	7	9
TEB	1, 6   1	583	5,066	5,264	10,330	6	6
TEB	24   24	125	1,654	1,738	3,392	1	2

third column shows the total number of hours the particular configuration was in use during the year. The fourth, fifth and the sixth columns present the total number of arrivals, departures and their sum, for the corresponding configurations, respectively. Finally, the seventh and the eighth columns list the percentages based on the total number of hours the airport was in operation and the total number of operations during the year. These percentages were computed by removing data corresponding to when the airport was closed or the configuration information was absent.

Next, the configurations at JFK, EWR, LGA and TEB were considered together at every hour of every day in 2017 to determine the most frequently used combinations of configurations. These combinations are summarized in Table 2. Comparing these tables, one observes that the top five configurations at JFK, EWR, LGA and TEB in Table 1 are

**Table 2 Top five most frequently used combinations of JFK, EWR, LGA and TEB configurations**

Configuration	Hours	Arrivals	Departures	Operations	% Hours	% Operations
JFK- 31L, 31R   31L; EWR- 4R   4L; LGA- 31   4; TEB- 6   1	410	38,998	38,822	77,820	5	5
JFK- 31L, 31R   31L; EWR-22L   22R; LGA- 22   13; TEB- 19   24	290	25,237	26,831	52,068	3	3
JFK- 31L, 31R   31L; EWR- 22L   22R; LGA- 22   31; TEB- 19   24	279	27,218	26,888	54,106	3	4
JFK- 31L, 31R   31L; EWR- 4R   4L; LGA- 31   4; TEB- 1, 6   1	264	25,901	26,779	52,680	3	3
JFK- 31L, 31R   31L; EWR- 4R   4L; LGA- 4   13; TEB- 6   1	264	24,727	25,933	50,660	3	3

used 74%, 84%, 86% and 96% of the time (based on hours), respectively, while the top five combinations of the configurations in Table 2 are only used 17% of the time. Of the 687 unique combinations of the configurations observed in the 2017 data, ten were used  $\geq 2\%$  of the time, 12 were used 1% of the time and the remaining 665 were used less than 1% of the time. The top ten and the top 22 combinations were used 28% and 40% of the time, respectively. It also turned out that the top five configurations listed in Table 2 consist of combinations of the top five most frequently used configurations at the four airports listed in Table 1. The configuration in the second row in Table 2 is composed of the most frequently used configuration at JFK, EWR, LGA and TEB in Table 1.

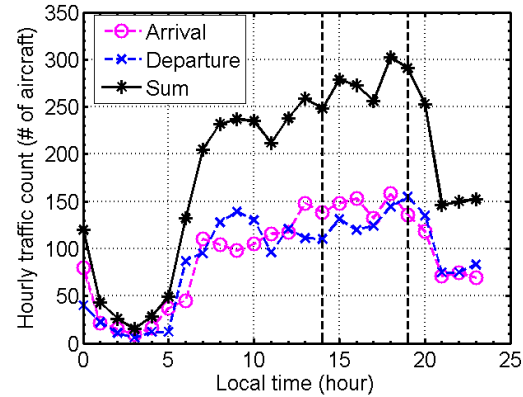
Finally, the sum of the operations at JFK, EWR, LGA and TEB were computed for every hour of every day in 2017 and the configuration data were sorted in non-increasing order of the sum of operations to identify dates and times with large number of operations. These results are summarized in Table 3. While archived traffic data do exist

**Table 3 Top five hours with most operations at JFK, EWR, LGA and TEB taken together**

Configuration	Arrivals	Departures	Operations	Local Time	UTC	Date
JFK- 13L   13R; EWR- 22L   22R; LGA- 22   13; TEB- 19   24	130	183	313	18	22	9/13/2017
JFK- 13L, 22L   13R; EWR- 4R, 11   4L; LGA- 22   13; TEB- 6   1	152	154	306	18	22	6/7/2017
JFK- 31L, 31R   31L; EWR- 22L   22R; LGA- 22   13; TEB- 19   24	158	144	302	18	22	5/23/2017
JFK- 13L, 22L   13R; EWR- 11, 22L   22R; LGA- 22   13; TEB-19   24	146	155	301	17	22	11/21/2017
JFK- 13L, 22L   13R; EWR- 11, 22L   22R; LGA- 22   13; TEB- 19   24	147	151	298	18	22	8/10/2017

for all the 2017 days in raw form in the ATM-data-warehouse, data up to May 2017 have been processed to remove duplicate records associated with the same flight. Given the availability of processed data, 5/23/2017 traffic data were chosen for generating the MACS scenario for simulating arrival traffic to JFK, EWR, LGA and TEB using the two-step method. Figure 1 shows the sum of traffic counts at each hour (local time), determined by summing the hourly traffic counts at JFK, EWR, LGA and TEB derived from 5/23/2017 ASPM data. The sum of the arrival counts, sum of the departure counts and the sum of the total counts are shown as three different graphs marked with ‘o’, ‘x’ and ‘\*’ symbols, respectively.

The two vertical dashed lines at local hour 14 (18 UTC) and 19 (23 UTC) mark the six-hour period chosen for analysis; it includes 22 UTC, the hour with the most number of operations on that day (see Table 3). This interval was also chosen because it had balanced operations (ratio of arrivals to departures close to one) with 865 arrivals and 784 departures. The MACS scenario was generated using the 5/23/2017 air traffic data archived in the ATM-data-warehouse to ensure that the traffic simulation would result in most flights arriving at JFK, EWR, LGA and TEB within this six-hour period. The MACS scenario generation is described next.



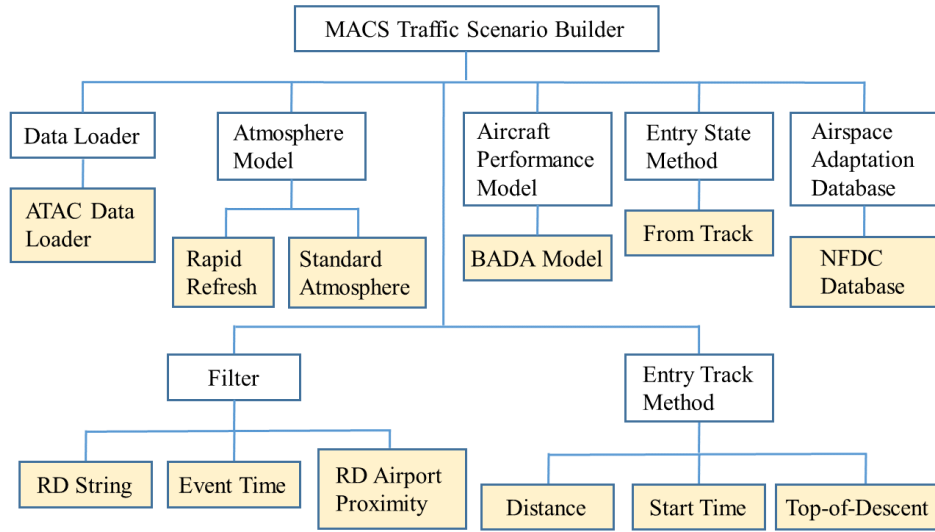
**Fig. 1 Hourly sum of traffic counts at JFK, EWR, LGA and TEB on 5/23/2017 based on ASPM data.**

#### IV. Step 1: Automated Scenario Generation Process

The automated scenario generation process in ATMTB is initiated by dragging and dropping predefined *blocks* (user defined and testbed native), specifying the block properties and linking the blocks graphically using the Simulation Architect, where the Simulation Architect application is launched from the web frontend as described in Ref. 6. While the details are available in Ref. 6, the description below is included as background for the second step of the two-step scenario generation process.

MACS scenarios are generated by processing recorded air traffic data archived in the ATM-data-warehouse, which is a platform for collecting, archiving, processing, querying and retrieving ATM data. Processed data derived from FAA’s System-Wide Information Management (SWIM) data are available in the ATM-data-warehouse in three different types of files- Reduced Record (RD), Event Data (EV) and Integrated Flight Format (IFF). RD files contain a single record for each flight, where the record contains information such as the reference time, unique key, aircraft ID, aircraft type, beacon code, airline, origin (airport or Fix-Radial-Distance (FRD)), takeoff/landing runway, destination airport, top-of-climb/top-of-descent time, runway threshold arrival/departure time, flight-plan (including route) data, and sector/center transition list. EV files contain multiple records for events related to each flight such as reference time, unique key, aircraft ID, aircraft type, event time and event type. MACS scenario generation currently processes takeoff/landing and crossing events, which includes sector, center and TRACON crossings. IFF files contain multiple records for each flight, where the records contain all flight-plans including amended flight-plans and track-data.

Figure 2 shows the various inputs that have to be specified for MACS scenario generation and the choices associated with them. Data contained in the RD and EV records are especially useful for filtering the traffic data for building the scenarios. The IFF data are useful for augmenting the traffic data derived from RD and EV records. Three types of filters are currently available. *RD String Filters* are used for selecting records from RD files by matching specified strings to those in the records. The user selects the desired filter from the list of filters and provides a list of strings appropriate for the selected filter. For example, airport code such as KEWR is a string that is compatible with the *Arrival/Departure Airport* filter. Similarly, 22L is an appropriate string in the list of strings with the *Landing Runway* filter option. The *RD Airport Proximity* filter is used for selecting flights to/from airports either inside or outside the specified region by processing RD records. The filter is set up by selecting from a list of options related to the properties like *Reference Location* and *Reference Distance* and inputting the values needed by the properties. Finally, *Event Time Filter* uses EV records to select flights. The Event Time Filter block has *Event Type*, *Minimum Value*, *Maximum Value* and *Include/Exclude* properties. Examples of Event Type are *Landing*, *Takeoff*, *Top-of-Climb* and *Top-of-Descent*. Minimum and Maximum Values are day (year-month-date) and UTC time (hour-minute-



**Fig. 2 Inputs and associated options for MACS scenario generation.**

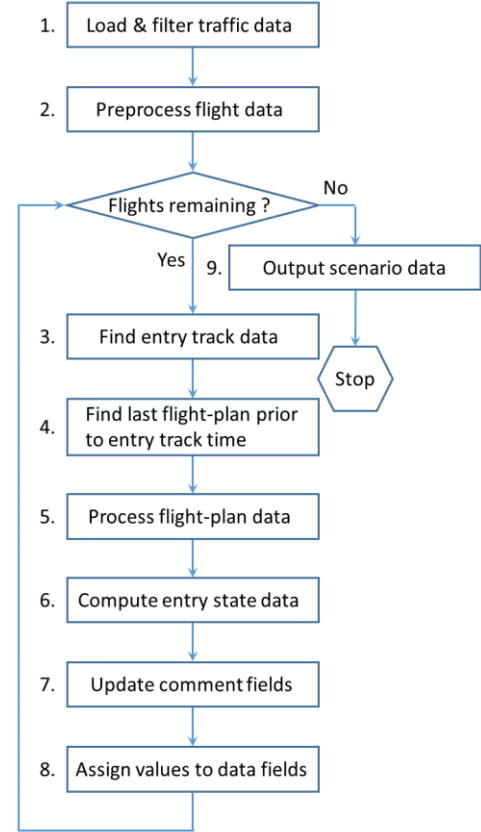
second)). The Include/Exclude property option specifies whether the flight events within the specified time interval are to be included or excluded.

In addition to the selection of data specified using filter blocks on the Simulation Architect, *Entry Track Method*, *Entry State Method*, *Aircraft Performance Model*, *Airspace Adaptation Database* and *Atmosphere Model* have to be specified. Three options for the Entry Track Method relevant to MACS scenario generation are: *Distance*, *Start Time* and *Top-of-Descent*. Target Airport ID and Distance from the Airport are the two parameters of the Distance block. Starting locations of the selected flights are chosen to be inside/outside the circular region defined by these two parameters. Start Time block enables the user to input the desired time past the simulation start time for selecting the starting position. For example, if the desired time is 30 minutes, the position of the flight at or just after when the simulation time is 30 minutes past the simulation start time would be chosen as the starting position. The Top-of-Descent block allows the user to specify a time with respect to top-of-descent for selecting the initial position of the flight. A value of -5 minutes would result in the selection of the position five minutes (or slightly more because track-data might not be available exactly at 5 minutes) prior to the time the flight reaches the top-of-descent point. At present, there is a single option associated with each of the other inputs. The only option available for the Entry State Method is *From Track*. Inclusion of the From Track block tells the scenario generator to use actual track data and the Mach transition altitude, determined using the BADA aircraft performance model and the specified atmospheric model, to determine the state of the flight such as altitude, heading, calibrated airspeed and Mach number at the entry location. The only option for Aircraft Performance Model is *BADA Model* block, and for Airspace Adaptation Database is *National Flight Data Center (NFDC) Database* block. Two options for the Atmosphere Model are *Rapid Refresh* block and *Standard Atmosphere* block.

The steps for MACS scenario generation are summarized in Fig. 3. The first step consists of loading RD, EV and IFF files from ATM-data-warehouse and filtering traffic data according to the filters specified on the Simulation Architect, and creating the flight data structure. The second step consists of assigning a BADA aircraft model in the flight data structure based on aircraft type and BADA *Synonym List*, and sorting the flight-plans of each flight by time. BADA Synonym List enables mapping of aircraft types that do not exist in the BADA database to the ones that exist in the database. The next step consists of finding the entry track data of the flights based on the simulation start time and the Entry Track Method specified on the Simulation Architect. Entry track data consist of time, latitude and longitude, altitude, groundspeed, course heading, Rate of Climb or Descent (ROCD) and sector ID of the *entry point*. The last flight-plan prior to entry track time is determined in the fourth-step. The fifth-step converts the flight-plan specified in terms of Departure Procedure (DP), airways, airway intersections, navigation aids, fixes, jet routes and Standard Terminal Arrival Route (STAR) into a sequence of latitude/longitude pairs. This step requires use of an airspace adaptation database like the NFDC Database. The parsed flight-plan route is augmented with the approach route from the STAR to the landing runway. Approach routes from STARs to landing runways are available to MACS as adaptation data. This step also determines the next waypoint, which could be an FRD, latitude/longitude location



or a named fix, along the flight-plan for connecting the entry point to the flight-plan, and builds the MACS route for the flight starting at the entry point and ending at the landing runway. Entry state data are determined in the sixth-step using track-data, which is specified by selecting the From Track block, the only available Entry State Method, on the Simulation Architect. Entry state data consist of true heading, calibrated airspeed, Mach, flight state (overflight, arrival or departure), and in-Mach or in-CAS mode at the entry point. MACS requires a target waypoint with speed and altitude constraints to be specified. The target waypoint is specified based on the flight state at the entry point. For flights in takeoff and climb phase at the entry point, the first waypoint after top-of-climb is chosen to be the target waypoint. If the flight is in cruise phase at the entry point, the next waypoint is chosen to be the target waypoint; if the next waypoint is beyond top-of-descent, the next waypoint with speed and altitude constraints on the approach route is chosen as the target waypoint; else, the airport is chosen as the target waypoint. If the flight is in descent phase at the entry point, the next waypoint with speed and altitude constraints on the approach route is chosen as the target waypoint. If the approach route is missing, the airport is chosen as the target waypoint. Data for several comment fields in the MACS scenario file are generated in the seventh-step. These data are useful for debugging and analysis. Values for all the data fields specified in the header of the version of MACS being used are assigned in the eighth-step based on the computations done in the earlier steps. The scenario data are output in a file in the last-step shown in Fig. 3.



**Fig. 3. Summary of MACS scenario generation steps.**

## V. Step 2: Scenario Refinement

The automated scenario generation process described in the previous section was employed to create a MACS scenario file for arrivals to JFK, EWR, LGA and TEB using 5/23/2017 traffic data archived in ATM data warehouse. All flights, including the ones with flight plans and the ones without flight plans, landing during the 18 through 23 UTC interval were considered for inclusion in the scenario. For flights without flight plans, the track data were processed to create flight-plans by specifying their route as a sequence of latitude-longitude pairs from the starting location to the destination airport. Cruise altitude and cruise speed were assigned based on the maximum altitude and maximum groundspeed seen in the track-data of these flights. The simulation start-time for creating the MACS scenario was chosen to be 12 UTC, which is six-hours prior to 18 UTC, to ensure that all flights in the scenario are able to land within the 18 UTC to 23 UTC time interval. The automated scenario generation process with these traffic data and scenario parameters resulted in the MACS scenario file with 808 flights, 57 fewer arrivals compared to 865 reported in ASPM.

Starting with the scenario file with 808 flights, the scenario refinement steps in Fig. 4 were employed to improve the data quality and to adjust the scenario for meeting experiment requirements. Flights with route-length of less than 20 nautical-miles eliminated in the first step reduced the number of flights by 17. In the second step, none of the flights were removed by filtering based on cruise speed because the minimum speed of 127 knots and the maximum speed of 571 knots are reasonable. Next, in step 3, filtering based on cruise altitude of more than 600 feet eliminated 22 additional flights whose cruise altitude was zero, most likely because of missing altitude information. VFR aircraft without a transponder with altitude reporting capability are not required to provide altitude reports automatically to air traffic control. The possibility of assigning a reasonable cruise altitude based on the performance characteristics of the type of aircraft and the length of route could be investigated in the future.

In step 4, the entry time filter was implemented to remove flights with an entry time of 30 minutes past the simulation start-time as suggested in Ref. 6. This step did not filter any flights because the earliest entry time in this dataset was one-hour and 28-minutes.

The internal to external flights ratio filter (step 5) is designed to eliminate a number of shorter and longer flights from the flights in the scenario file to achieve the desired ratio. Flights are categorized as internal flights if the length of route is less than a prescribed threshold and external otherwise. A threshold value of 400 nautical-miles was used in this study. Of the 769 remaining flights at the end of step-four in Fig. 4, 238 were categorized as internal flights and 531 as external flights with the resulting internal to external flights ratio of 0.45. Let, the desired ratio be  $r$ ,  $x_1$  be the number of external flights in the dataset,  $x_2$  be the number of internal flights in the dataset,  $x_{1s}$  be the number of selected external flights and  $x_{2s}$  be the number of selected internal flights. The procedure for selecting  $x_{1s}$  and  $x_{2s}$  such that

$$\frac{x_{2s}}{x_{1s}} = r \quad (1)$$

is given by:

$$x_{1s} = x_1 \text{ and } x_{2s} = \lfloor rx_1 \rfloor \text{ if } r \leq \frac{x_2}{x_1} \quad (2)$$

and

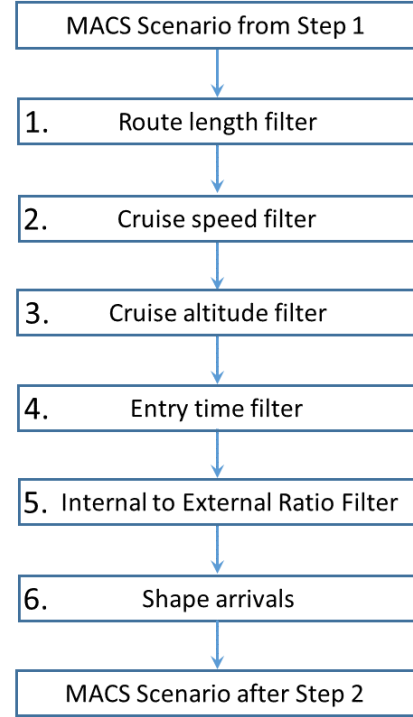
$$x_{1s} = \left\lfloor \frac{x_2}{r} \right\rfloor \text{ and } x_{2s} = x_2 \text{ if } r > \frac{x_2}{x_1} \quad (3)$$

These two solutions can be written together as follows using the Iverson's notation:

$$\begin{aligned} x_{1s} &= [rx_1 \leq x_2]x_1 + (1 - [rx_1 \leq x_2])\left\lfloor \frac{x_2}{r} \right\rfloor \\ x_{2s} &= [rx_1 \leq x_2]\lfloor rx_1 \rfloor + (1 - [rx_1 \leq x_2])x_2 \end{aligned} \quad (4)$$

where the logical expressions within the square-brackets mean a value of one or zero depending on whether they are true or false. Table 4 shows the number of internal and external flights for different values of  $r$ . Observe that a large value of  $r$  like 300 results in only internal flights to be selected; with zero external flights selected, the desired internal to external ratio is  $\infty$ . Finally, the internal and external flights are selected by first sorting the lists of internal flights and external flights in non-increasing order of route length, and then picking the required number starting from the top of the two lists. Alternatively, the required numbers can be selected randomly from the two unordered lists of internal and external flights.

Step 5 for selecting internal and external flights based on their desired ratio, though implemented, was not applied; all 769 flights in the scenario were accepted. Figure 5 shows the hourly arrival traffic count considering arrivals to



**Fig. 4 Scenario refinement steps.**



JFK, EWR, LGA and TEB together, as a function of UTC. The figure also shows the predicted hourly arrival count based on the estimate of landing times

$$\hat{t}_L = t_E + \frac{l_R}{V_{Avg.}} \quad (5)$$

where  $\hat{t}_L$  is the predicted landing time,  $t_E$  is the entry time (takeoff time for aircraft on the ground),  $l_R$  is the route length and  $V_{Avg.}$  is the average cruise groundspeed, which is determined by averaging the actual cruise speed derived from track-data within the top-of-climb and top-of-descent interval. Predicted landing rate comparison with the actual landing rate is useful for a sanity check. Hourly arrival traffic counts shown in Fig. 5 were determined as the number of arrival flights in the hourly window, continuously shifted temporally at a five-minute interval. The figure suggests that MACS scenario with average cruise groundspeed would result in a scenario that would reasonably replicate the actual landing rate.

Finally, the last requirement of being able to shape the arrival traffic (Step 6) from Ref. 6 was achieved by employing a first-come first-served scheduler. Let the proposed time of arrival provided by the aircraft operator be given by  $t_p$ ; this could also be the estimated time of arrival obtained using Eq. (5). Let the scheduled time of arrival determined by the first-come first-served scheduler be given by  $t_s$ . After sorting the flights in non-decreasing order by  $t_p$ , the scheduled time of arrival is determined using the following recursion equation:

$$t_s(i) = t_s(i-1) + \max(\Delta t, t_p(i) - t_s(i-1)) \quad (6)$$

starting with  $t_s(1) = t_p(1)$ . For the desired hourly arrival

count of  $\dot{n}$  aircraft, minimum temporal separation,  $\Delta t$ , between successive aircraft is then  $60/\dot{n}$  minutes. Equation (6) states that if the following aircraft  $i$  is temporally separated from the leading aircraft  $i-1$  by more than  $\Delta t$ , set the scheduled arrival time of the following aircraft to its originally proposed time of arrival; if not separated, add  $\Delta t$  to the scheduled arrival time of the leading aircraft and assign it to the following aircraft as the scheduled time of arrival. Thus, this scheduler only delays aircraft. The resulting delay is given as

$$\delta(i) = (\Delta t - t_p(i) + t_s(i-1)) \left[ t_p(i) - t_s(i-1) < \Delta t \right] \quad (7)$$

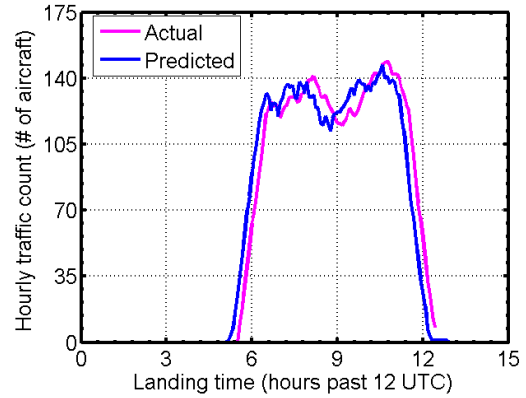
where the logical expression inside the square-brackets means a value of one when true and zero otherwise. Scheduling results were generated by imposing an hourly arrival capacity constraint of 100 aircraft. Figure 6 shows the original unconstrained hourly arrival traffic counts as a function of time (shown earlier in Fig. 5) and the traffic counts resulting from scheduling traffic using Eq. (6) to meet the specified capacity constraint. The average delay was found to be 49 minutes, maximum delay was one-hour and 44 minutes, and total delay was 634 hours.

To achieve the desired arrival schedule of traffic simulated by MACS, the entry time of the flights have to be adjusted. Because the flight time is given as:

$$t_f = t_p - t_E \quad (8)$$

**Table 4 Number of internal and external flights based on  $r$  values**

$r$	Internal	External
0	0	531
0.25	132	531
0.5	238	476
0.75	238	317
1	238	238
300	238	0



**Fig. 5 Actual and predicted hourly arrival traffic counts at JFK, EWR, LGA and TEB taken together.**

where  $t_E$  is the original entry time, the new scheduled entry time,  $t_{sE}$ , can be determined as:

$$t_{sE} = t_s - t_f = t_s - t_p + t_E \quad (9)$$

Figures 7 and 8 show the histograms of the original entry times and the new scheduled entry times. Observe the reduction in the number of flights in the bins and the spreading of the flights and entry times to beyond 12 hours past 12 UTC in Fig. 8.

The scheduling procedure described in this section is useful both for IDM HITL and fast-time MACS based simulations. This procedure can be used for allocating ground delay in response to airport capacity constraint forecast to achieve strategic traffic flow management objectives. The same procedure can then be used to allocate airborne delay to flights in the terminal area in response to actual airport capacity constraint to achieve tactical traffic flow management objectives. A slightly modified version of the scheduler in Eq. (6),

$$t_s(i) = t_s(i-1) + \Delta t \quad (10)$$

where the delay,  $\Delta t$ , could be changing as a function of time, can be used for increasing the arrival traffic demand beyond the capacity of the airport as was needed for the manually modified HITL scenario in Ref. 6; the HITL scenario in Ref. 6 achieved demand exceeding capacity by squeezing six-hours of traffic into five hours.

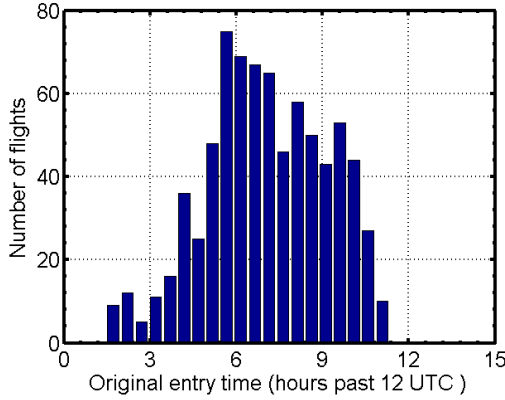


Fig. 7 Histogram of the original entry times.

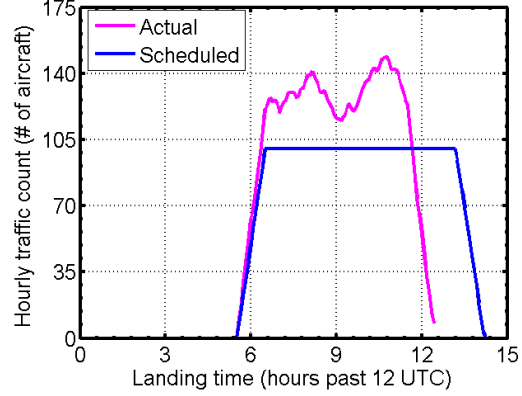


Fig. 6. Actual and scheduled hourly arrival traffic counts at JFK, EWR, LGA and TEB taken together.

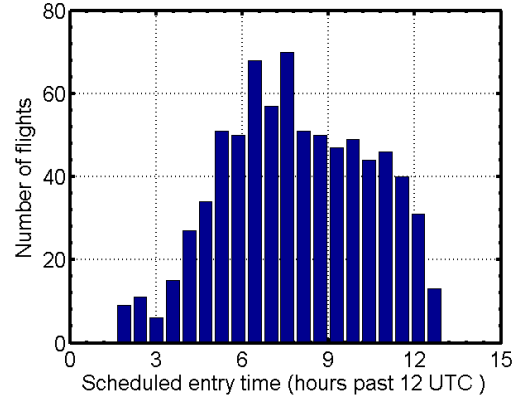


Fig. 8 Histogram of the scheduled entry times.

## VI. Conclusions

The two-step procedure for automated scenario generation for Multi-Aircraft Control System based traffic simulation was described. The first step utilized the scenario generation process currently being used by the Air Traffic Management Testbed in development at NASA Ames Research Center. The second step, which implemented refinements to the scenario output from the first step, for meeting the objectives of the Human-in-the-Loop experiments and fast-time simulations, was also described. Flights were filtered in the second step based on route length, cruise speed, cruise altitude, entry time and the ratio of internal to external flights. The procedure for selecting internal and external flights was described. Finally, first-come first-served schedulers were described for curtailing arrival traffic demand to meet the airport arrival capacity constraints, and to increase arrival traffic demand over the

airport capacity to meet Human-in-the-Loop experiment objectives. To determine the most frequently used runway configurations and the ones used during the busiest periods in terms of the number of operations (sum of arrivals and departures), runway configurations used during every hour of every day in 2017 and the associated numbers of arrivals and departure were obtained from the FAA's Aviation System Performance Metrics database. Results of analysis were presented in tables to summarize top five most frequently used configurations at John F. Kennedy, Newark Liberty, LaGuardia and Teterboro airports both individually and together. This analysis led to the identification of 5/23/2017 as a busy traffic day on which to base the scenario generation. Flights with and without flight plans in the 5/23/2017 traffic data were processed using the two-step procedure and the scheduling procedure to generate the results. These results show that the automated procedures discussed in the paper can be used to generate traffic scenarios that meet the requirements of Human-in-the-Loop experiments and fast-time simulations for evaluation of air traffic concepts. The automated process can replace the tedious manual scenario generation process; it is less error prone and makes it possible to generate large number of scenarios needed for Monte-Carlo evaluations, which is very difficult to achieve with the manual process.

### Acknowledgements

Authors are grateful to Dr. Min Xue, Dr. Antony Evans, Confesor Santiago, William Chan and Katharine Lee for reviewing the paper and providing feedback. Discussions with Dr. Robert Windhorst are gratefully acknowledged.

### References

- <sup>1</sup>Prevot, T., et. al., "Human-in-the-Loop Evaluation of NextGen Concepts in the Airspace Operations Laboratory," AIAA 2010-7609, *AIAA Modeling and Simulation Technologies Conference*, Toronto, Ontario, Canada, August 2-5, 2010.
- <sup>2</sup>Gong, C., McNally, D., and Lee, C. H., "Dynamic Arrival Routes: A Trajectory-Based Weather Avoidance System for Merging Arrivals and Metering," AIAA 2015-3394, *15<sup>th</sup> AIAA Aviation Technology, Integration, and Operations Conference*, Dallas, Texas, June 22-26, 2015.
- <sup>3</sup>Smith, N. M., et. al., "Integrated Demand Management: Coordinating Strategic and Tactical Flow Scheduling Operations," *16<sup>th</sup> AIAA Aviation Technology, Integration, and Operations Conference*, Washington, DC, 2016.
- <sup>4</sup>Ging, A., et. al., "Airspace Technology Demonstration 2 (ATD-2) Technology Description Document," NASA Technical Memorandum: NASA/TM-2018-219767, NASA Center for AeroSpace Information, 7115 Standard Drive 5301, Hanover, MD 21076-1320. URL: <https://www.aviationsystemsdivision.arc.nasa.gov/publications/2018/NASA-TM-2018-219767.pdf> [cited 5/10/2018].
- <sup>5</sup>Chan, W. N., "ATM-X: Air Traffic Management - eXploration," *Partnership Workshop*, NASA Ames Research Center, Moffett Field, California, April, 2018. URL: <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20180002413.pdf>. [cited 5/10/2018].
- <sup>6</sup>Chatterji, G. B., et. al., "Automated Scenario Generation for Human-in-the-Loop Simulations," *AIAA Modeling and Simulation Technologies Conference*, Atlanta, GA, June 25-29, 2018.
- <sup>7</sup>Moehlenbrink, C., et. al., "Evaluation of Integrated Demand Management looking into Strategic & Tactical Flow Management," *12<sup>th</sup> USA/Europe Air Traffic Management Research and Development Seminar*, Seattle, Washington, June 27-30, 2017. URL: [http://www.atmseminarus.org/seminarContent/seminar12/papers/12th\\_ATM\\_RD\\_Seminar\\_paper\\_51.pdf](http://www.atmseminarus.org/seminarContent/seminar12/papers/12th_ATM_RD_Seminar_paper_51.pdf) [cited 5/10/2018].
- <sup>8</sup>Chan, W. N., et. al., "Overview of NASA's Air Traffic Management – eXploration (ATM-X) Project," *AIAA Aviation Technology, Integration, and Operations Conference*, Atlanta, GA, June 25-29, 2018.
- <sup>9</sup>Yoo, H., et. al., "Required Time of Arrival as a Control Mechanism to Mitigate Uncertainty in Arrival Traffic Management," *35<sup>th</sup> IEEE Digital Avionics Systems Conference*, Sacramento, California, 2016.
- <sup>10</sup>Yoo, H., et al., "Integrated Demand Management: Minimizing Unanticipated Excessive Departure Delay while Ensuring Fairness from a Traffic Management Initiative," *17th AIAA Aviation Technology, Integration, and Operations Conference*, 5-9 June 2017.
- <sup>11</sup>Arneson, H., Evans, A. D., Kulkarni, D., Lee, P., Li, J., Wei, M. Y., "Using an Automated Air Traffic Simulation Capability for a Parametric Study in Traffic Flow Management," *18th AIAA Aviation Technology, Integration, and Operations Conference*, Atlanta, GA, June 24-28, 2018.
- <sup>12</sup>Eurocontrol Experimental Centre, "User Manual for the Base of Aircraft Data (BADA) Revision 3.6," EEC Note No. 10/04, Eurocontrol Experimental Centre Publications Office, B.P. 15, 91222 - Bretigny-sur-orge, France, July, 2004.